

## WHAT IS CLAIMED IS:

- 1           1.     A method of segmenting video input characterized by a time series  
2 of video frames of observable pixel data, comprising:  
3           maintaining one or more pixel-level historical models of spatially local  
4 pixel observations;  
5           segmenting pixels into two or more labeled groups based at least in part  
6 upon comparison of pixel-level video input with the one or more pixel-level  
7 historical models; and  
8           updating the one or more pixel-level historical models based at least in part  
9 upon one or more feedback maps identifying pixels respectively segmented into  
10 the one or more labeled groups in conformity with a spatially non-local  
11 segmentation model.
- 1           2.     The method of claim 1, wherein a history of pixel observations at  
2 each pixel is modeled by a mixture of Gaussian distributions.
- 1           3.     The method of claim 1, wherein pixels are segmented into a  
2 background group and a foreground group.
- 1           4.     The method of claim 1, wherein a feedback map identifies pixels  
2 segmented correctly according to a spatially non-local segmentation model.
- 1           5.     The method of claim 4, wherein the spatially non-local segmentation  
2 model defines spatially non-local observation characteristics of pixels belonging to  
3 one of the labeled groups.
- 1           6.     The method of claim 4, wherein a pixel-level historical model is not  
2 updated at pixels identified as being correctly segmented.
- 1           7.     The method of claim 1, wherein a feedback map identifies pixels  
2 segmented incorrectly according to a spatially non-local segmentation model.
- 1           8.     The method of claim 7, wherein the spatially non-local segmentation  
2 model defines spatially non-local observation characteristics of pixels that should  
3 have been excluded from one of the labeled groups.

1           9.     The method of claim 7, wherein a pixel-level historical model is  
2     updated at pixels identified as being incorrectly segmented.

1           10.    The method of claim 9, wherein updating the pixel-level historical  
2     model at pixels identified as being incorrectly segmented comprises maintaining a  
3     per-pixel inclusion error model of pixel observations associated with occurrences  
4     of incorrect segmentation labeling.

1           11.    The method of claim 10, wherein components of the per-pixel  
2     inclusion error model corresponding to individual pixels include respective  
3     mixtures of Gaussian distributions.

1           12.    The method of claim 11, wherein updating a pixel-level historical  
2     model comprises merging a per-pixel historical model and a per-pixel inclusion  
3     error model.

1           13.    The method of claim 1, further comprising generating the feedback  
2     maps based at least in part upon analysis of spatially non-local video frame  
3     features.

1           14.    The method of claim 13, wherein the feedback maps are generated  
2     based at least in part upon one or more of an image region analysis, a frame-wide  
3     image statistics analysis, or an analysis of the object or event content of the video  
4     frames.

1           15.    The method of claim 14, wherein one or more of the feedback maps  
2     are generated based at least in part upon depth information or stereo disparity  
3     information, or both.

1           16.    The method of claim 1, further comprising generating one or more  
2     confidence maps associating pixels with respective measures of segmentation  
3     accuracy.

1           17.    The method of claim 16, further comprising merging multiple  
2     confidence maps to produce a merged confidence map.

1           18.    The method of claim 17, wherein the measures of segmentation  
2 accuracy are real numbers, and the step of merging multiple confidence maps  
3 comprises adding the multiple segmentation accuracy measures respectively  
4 associated with each pixel.

1           19.    The method of claim 17, further comprising thresholding the merged  
2 confidence map to produce one or more feedback maps.

1           20.    The method of claim 16, wherein each of the confidence maps is  
2 generated based at least in part upon one or more of an image region analysis, a  
3 frame-wide image statistics analysis, or an analysis of the object or event content  
4 of the video frames.

1           21.    The method of claim 20, wherein a pixel-level historical model  
2 includes a mixture of Gaussian distributions of pixel observations.

1           22.    The method of claim 1, wherein one or more pixel-level historical  
2 models incorporate per pixel depth information or stereo disparity information, or  
3 both.

1           23.    The method of claim 1, wherein pixels are segmented based at least  
2 in part upon depth information or stereo disparity information, or both.

1           24.    The method of claim 1, wherein one or more feedback maps are  
2 generated by one or more of a person detector and tracker module, a rapid  
3 illumination change detector module, a camera gain change detector module, or a  
4 sudden camera motion detector module.

1           25.    A system for segmenting video input characterized by a time series  
2 of video frames of observable pixel data, comprising one or more processing  
3 modules operable to:

4           maintain one or more pixel-level historical models of spatially local pixel  
5 observations;

6           segment pixels into two or more labeled groups based at least in part upon  
7 comparison of pixel-level video input with the one or more pixel-level historical  
8 models; and

9           update the one or more pixel-level historical models based at least in part  
10   upon one or more feedback maps identifying pixels respectively segmented into  
11   the one or more labeled groups in conformity with a spatially non-local  
12   segmentation model.

1           26.    The system of claim 25, wherein a history of pixel observations at  
2   each pixel is modeled by a mixture of Gaussian distributions.

1           27.    The system of claim 25, wherein pixels are segmented into a  
2   background group and a foreground group.

1           28.    The system of claim 25, wherein a feedback map identifies pixels  
2   segmented correctly according to a spatially non-local segmentation model.

1           29.    The system of claim 28, wherein the spatially non-local  
2   segmentation model defines spatially non-local observation characteristics of  
3   pixels belonging to one of the labeled groups.

1           30.    The system of claim 28, wherein a pixel-level historical model is not  
2   updated at pixels identified as being correctly segmented.

1           31.    The system of claim 25, wherein a feedback map identifies pixels  
2   segmented incorrectly according to a spatially non-local segmentation model.

1           32.    The system of claim 31, wherein the spatially non-local  
2   segmentation model defines spatially non-local observation characteristics of  
3   pixels that should have been excluded from one of the labeled groups.

1           33.    The system of claim 31, wherein a pixel-level historical model is  
2   updated at pixels identified as being incorrectly segmented.

1           34.    The system of claim 33, wherein the pixel-level historical model is  
2   updated at pixels identified as being incorrectly segmented by maintaining a per-  
3   pixel inclusion error model of pixel observations associated with occurrences of  
4   incorrect segmentation.



3 frame-wide image statistics analysis, or an analysis of the object or event content  
4 of the video frames.

1 44. The system of claim 43, wherein a pixel-level historical model  
2 includes a mixture of Gaussian distributions of pixel observations.

1 45. A computer program for segmenting video input characterized by a  
2 time series of video frames of observable pixel data, the computer program  
3 residing on a computer-readable medium and comprising computer-readable  
4 instructions for causing a computer to:

5 maintain one or more pixel-level historical models of spatially local pixel  
6 observations;

7 segment pixels into two or more labeled groups based at least in part upon  
8 comparison of pixel-level video input with the one or more pixel-level historical  
9 models; and

10 update the one or more pixel-level historical models based at least in part  
11 upon feedback maps identifying pixels respectively segmented into the one or  
12 more labeled groups in conformity with a spatially non-local segmentation model.